DOCUMENT RESUME

ED 105 539 CS 501 054

AUTHOR Shellen, Wesley N.
TITLE Measurement of Listening Achievement: The Role of Perceived Interest and Extrinsic Incentives.
PUB DATE Apr 75
NOTE 19p.; Paper presented at the Annual Meeting of the International Communication Association (Chicago, April, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS College Students; *Educational Research; Higher Education; *Interest Research; Interests; *Listening Comprehension; Listening Habits; *Listening Skills; *Measurement

ABSTRACT

This experiment was designed to test the influence of selected variables characteristic of both normal listening and listening under test conditions to determine whether test incentives negate or interact with the normal listening process. Public speaking classes at the University of Montana were asked as part of their regular classwork to rate the interestingness of recorded messages played to them over earphones. The key implication of the study is that test incentives serve to increase listening achievement scores of students but do not negate or interact with perceived interest, a variable related to the normal listening process. Therefore, listening achievement scores obtained under test conditions may be interpreted as being more representative of normal listening behavior than was previously believed. (Author/RB)

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Wesley N. Shellen

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

ED105539

MEASUREMENT OF LISTENING ACHIEVEMENT: THE ROLE OF PERCEIVED INTEREST AND EXTRINSIC INCENTIVES

Wesley N. Shellen

Critics of listening research distinguish between so-called "normal" listening versus listening under test conditions (Kelly, 1963). Consequently, listening tests are said to measure achievement under exaggerated incentive conditions not characteristic of the normal listening situation. Weaver (1972b) explained that listening achievement under normal conditions will be governed by (1) the listener's willingness to listen and (2) the listener's ability or capacity to listen. The willingness to listen is affected by the listener's interest in or agreement with a message stimulus. The ability or capacity to listen is affected by the difficulty of the message or the rate of speed at which the message is presented.

Kelly (1963) discovered that subjects' scores on a surprise listening test following a lecture correlated significantly with several personality measures but not with their scores on a test of mental ability. When the subjects were aware that they were to be tested, however, their scores correlated with mental ability but not with the personality factors. This finding led him to the conclusion that listening under normal conditions was somehow qualitatively different from listening under test conditions. Thus he argued that the demand characteristics within the instructional set of listening test conditions preclude the influence of the listener's willingness to listen, consequently measuring only the listener's ability or capacity to listen.

CS 501 054

When Spearritt (1962) conducted his factor analysis of listening tests, he was concerned only with the measurement of listening ability, not habits or willingness. His review of the literature led him to an assumption which became one of the central hypotheses of the present experiment. Spearritt reasoned that:

> If adequate motivation among the subjects is secured in the testing situation, the effect of differential interest in the material presented is likely to be reduced or eliminated [p. 8].

Spearritt cited two sources of evidence in support of the assumption. The first was a study showing that listeners retain more material if they know they are to be tested (Knower, Phillips, & Koeppel, 1945). But since no ratings of interest were obtained in this study, it allowed no comparison of the effects of awareness of testing between interested and disinterested listeners. In Spearritt's own study, all subjects were award that they were in a test situation, so no comparison group of unmotivated listeners was available to test the assumption that interest would affect their listening more than that of the motivated subjects. The most direct evidence cited by Spearritt was from an article by Brown (1955), which was a report of an informal classroom exercise, not a controlled experiment.

The present experiment was designed to test the influence of selected variables characteristic of both normal listening and listening under test conditions to determine whether test incentives actually negate or interact with listener's willingness to listen.

Much prior research has established the generalization

that listeners who are interested in a message tend to outscore uninterested listeners on tests over the message contents (Trenaman, 1951; Vernon, 1950; Brandon, 1956; Matter, 1968; Young, 1972; Weaver, 1972a). Livingston (1961) reported the only exception to this generalization. Further, some messages were found to appeal differently to the interests of male versus female subjects, thus accounting for a sex difference in listening performance (Spearritt, 1962; Klein, 1969). Listening achievement has not been found to relate to vocational and academic interests (Karraker, 1951; Heath, 1951, Marten, 1958). Messages differing in Human Interest as measured by the Flesch scale affected listening only when the experimental messages differed greatly in Human Interest (Allen, 1952; Cartier, 1955).

Various researchers have offered extrinsic incentives to motivate subjects to perform well on listening achievement tests. Sewell (1972) found that monetary incentives did not increase the listening achievement scores of college students. Academic class grades have generally been found more successful as incentives to improve listening achievement (Bohn & Frandsen, 1964; Goodyear, 1969). The incentives for the present experiment were chosen primarily because they duplicated the conditions described by Spearritt, Kelly, and others concerned with the distinction between listening behavior under normal audience conditions and testing conditions. The procedures established for the experiment were designed to appear as though they were a part of the planned schedule of the classes in which the subjects were enrolled.

Three major hypotheses were formulated for the study corresponding to the two-way analysis of variance design used:

1. Subjects who rate a message as interesting will score higher on an achievement test over its contents than subjects who rate the message as uninteresting.

2. Subjects listening to a message will score higher on a test over its content if (a) they are instructed prior to the message that they will be tested, and (b) if they are instructed that their test scores will apply toward a course grade.

3. There will be an ordinal interaction between expressed interest and the levels of test incentive. Specifically, increasing the level of test incentive will decrease the difference in test scores between subjects expressing high versus low interest in the message.

METHOD

Subjects

The subjects for the experiment were nonvolunteer undergraduate students enrolled in the Introduction to Public Speaking classes at the University of Montana during the fall quarter of 1972. One hundred and ninety-five students participated in the experiment. However, the main analyses were performed using only the data from 117 subjects who rated the experimental message on the upper and lower thirds of a six-point interest rating scale ranging from "quite interesting" to "quite uninteresting." Discarding of data from subjects rating the middle third of the scale was done to avoid the problem of regression effects

mentioned by Campbell and Stanley (1963) and to provide a more powerful test of the interest variable.

Experimental Messages and Test

The experimental messages consisted of two nonfiction prose selections, recorded on audio tape. The first message was an article by Armstrong (1971) exposing commercialized faith healing schemes. The purpose of the first message was to allow the subjects to become oriented to the experimental task, and to become familiar with the interest rating procedures. Also, no test was given over the content of this first selection to prevent students from expecting a test over the second message unless they were so instructed.

The second message, the one used for administering the actual experimental procedures, was from an article by Davis (1966), a rhetorical criticism of the famous evangelist Billy Sunday. Pilot data gathered prior to the experiment indicated that subjects' ratings of the interestingness of this message represented a full range of responses from the highest to the lowest ends of the interest rating scale. Further, the pilot data indicated that the interested subjects significantly outscored the uninterested subjects in their listening achievement test scores over the content of the message ($t = 2.17$; $p < .05$). The criteria resulting in the selection of this message for the experiment clearly biased the outcome of the analysis of the main effect due to interest. However, the literature is rich with instances of significant relationships between perceived interest and listening achievement, and the

experimenter considered it necessary to estimate the relationship in advance of the experiment in order to allow for a reasonable test of the interaction hypothesis between perceived interest and extrinsic incentives.

The listening achievement test used in this study was a twenty-item short answer test of the recall of facts explicitely presented in the experimental message. The test was assessed and revised through two pilot administrations. The total reliability of the revised listening achievement test over the contents of the Davis message was .86 as measured by the Kuder Richardson 20 formula. The item analysis of the revised test also showed that the point biserial correlations of all twenty items were significant. The difficulty indexes ranged from .17 to .90 with .58 being the mean difficulty index.

Experimental Procedures

The experiment took place in the communication recording laboratory at the University of Montana. The laboratory had twenty-one listening stations, each equipped with a stereo cassette recorder and earphones. The entire experimental treatment, including messages and instructions were recorded on individual cassettes.

The public speaking classes received the treatments in the laboratory during their normal class meeting time. When the subjects arrived at the laboratory they were given booklets containing the interest rating scales. Each subject was asked to take a seat at the booth to which he had been randomly assigned and was asked to put on the earphones. The recorders were pre-set to play

when the power from the main console was turned on. Since all treatments were administered in the same room by the same experimenter at the same time, this eliminated the need for any sort of double-blind procedures to control for expectancy effects.

The recorded instructions informed subjects that their confidential ratings of the interestingness of messages was needed for a listening training program to be included in future public speaking courses. Subjects then heard the first part of the Armstrong message and made initial ratings to express their level of interest. They then heard the rest of the message and made final interest ratings.

The first part of the Davis message, the actual experimental message, was then played to all subjects. This part was one minute and fifteen seconds long. Following the introduction, the message was interrupted as with the first selection and the listeners heard the same instructions asking for the initial rating. This set of initial interest ratings of the experimental message were the ratings used subsequently to block subjects into interested and uninterested groups for the experimental design. After the initial rating instructions the subjects received one of the three test incentive instructions. The use of initial ratings for the blocking was necessary to keep the interest ratings independent of effects caused by the incentive instructions. Trenaman (1967) justified the use of initial ratings for such experimental blocking by his finding that perceptions of interest in a message are rapidly formed by listeners and that initial ratings of interest correlate significantly with final interest ratings. A post check on this

assumption in the present experiment determined that subjects divided by initial ratings into groups of interested versus uninterested listeners also differed significantly in their final interest ratings after hearing the entire message ($F = 87.51$; $df = 1,111$).

The incentive instructions which followed the delay for initial interest ratings were of three types thus creating the three extrinsic incentive conditions for the experiment. One third of the subjects heard the same instructions that the message was about to continue which had preceded the remainder of the earlier message they heard. Thus, they were not warned that they would be tested following the message (no test incentives). The next third of the subjects additionally received a warning that they would be tested at the conclusion of the message but were told that their test scores would not apply toward their class grades (nongraded test incentives). The final third of the subjects were warned that they would be tested at the conclusion of the message and were told that their test scores would apply toward their class grades (graded test incentives).

The remainder of the experimental message was then played to the listeners. At the conclusion of the selection they were instructed to complete the final interest ratings. After completing the final rating, the subjects were asked to turn over their rating sheets and use the backside as an answer sheet for the test. The test was then administered on the tape with pauses after each question to allow time for subjects to write their answers.

To offset potential feelings of deception, all subjects

were later offered the choice of whether to apply the test scores to their class grades or not.

Design and Statistical Procedures

The data from the experiment were analyzed using a 2 x 3 randomized block design using the fixed effects linear model (Kirk, 1968). There were two levels in the interestingness factor and three levels in the extrinsic test incentive factor. The dependent variable consisted of the subjects' scores on the twenty-item listening achievement test over the content of the experimental message. The .05 level of significance was the criterion for rejecting the null hypothesis in each statistical test.

Since the design of the study used only data from subjects rating the message on the upper or lower thirds of the initial interest scale, the resulting cell frequencies were unequal and disproportional. Therefore, the analysis of variance was conducted using the least squares method which does not require equal or proportional numbers of subjects in each cell. Post-hoc pairwise multiple comparisons were made using the Newman-Keuls method for unequal replications (Kramer, 1956).

RESULTS

The design format and mean scores on the twenty-item listening achievement test for each group are presented in Table 1. The analysis of variance, presented in Table 2, showed that the overall F ratio for the interaction was not significant. The F ratio for the interestingness factor was significant, indicating

that subjects who rated the message as interesting scored significantly higher on the listening achievement test than subjects who rated the message as uninteresting.

TABLE 1

Experimental Design Format and Mean Scores on the Listening Achievement Test

Extrinsic Test Incentive Conditions	Interestingness Conditions: High Initial Interest	Interestingness Conditions: Low Initial Interest	Row Means
No Test Incentive	13.90 n = 20	9.65 n = 20	11.78
Nongraded Test Incentive	13.20 n = 24	10.67 n = 15	11.90
Graded Test Incentive	14.45 n = 20	12.83 n = 18	13.61
Column Means	13.85	11.01	

TABLE 2

Analysis of Variance of Scores on the Listening Achievement Test

Source	SS	df	MS	F
Interestingness	230.14	1	230.14	24.54*
Incentives	80.76	2	40.38	4.30*
Interaction	34.88	2	17.44	1.86
Within	1041.09	111	9.38	

* $p < .05$

The overall F ratio for the extrinsic test incentive factor was also significant. Since there were three levels of this factor, it was necessary to make post hoc multiple comparisons among the means of each group to determine which two or more groups differed significantly from one another. The results of the multiple comparisons are presented in Table 3.

TABLE 3

Pairwise Multiple Comparisons Between Mean Listening Achievement Scores for the Three Incentive Conditions

Group Comparisons	Newman-Keuls Multiple Range
No Test Incentives versus Nongraded Test Incentives	.81
No Test Incentives versus Graded Test Incentives	11.45*
Nongraded Test Incentives versus Graded Test Incentives	10.58*

* $p < .05$

The comparison analysis showed that the subjects who were warned of the test and told that they would receive a grade based on their achievement (graded test incentive) scored significantly higher than subjects who were either not aware they would be tested (no test incentive) or subjects who were aware they would be tested but were told that they would not be graded on their achievement

(nongraded test incentive). The mean score of the subjects who were not aware they would be tested did not differ significantly from the mean score of the subjects who were aware they would be tested but told that they would not be graded on their achievement.

DISCUSSION

The first hypothesis stated,

> Subjects who rate a message as interesting will score higher on an achievement test over its contents than subjects who rate the message as uninteresting.

The results of the analysis of variance showed that the null hypothesis regarding the relationship between interest and listening achievement could be rejected. Significantly higher listening test scores over the content of the experimental message were achieved by subjects who rated the message as interesting than subjects who rated it as uninteresting.

The extent to which this finding can be generalized to other messages should be qualified. The review of literature in which the relationship between interest and listening achievement was investigated found that the two variables were significantly related only for some messages. The experimental message for the present study was selected partly because the distribution of interest ratings ranged from both extremes of the scale, showing that listeners differed greatly in their perceptions of its interestingness. The experiment by Livingston (1961), however, used a message which was rated as either "interesting" or "highly

interesting" by all eighty of his subjects. None of the subjects rated the message as "boring." The correlation between interest ratings and scores on the listening achievement test used in Livingston's study was not significant. The conflicting results between the present study and the study by Linvingston might be explained statistically in that the narrow range of interest ratings obtained in Livingston's study might have provided a less reliable distinction between levels of perceived interest than the wide range of interest ratings obtained in the present study. The fact that none of Livingston's subjects perceived the message as boring may simply indicate that interest was not a variable in his study. If this explanation is accurate, the significant relationship between interest and listening achievement found in the present study might be generalized only to conditions where listeners differ greatly in their perceptions of the interestingness of a message.

The second hypothesis stated,

> Subjects listening to a message will score higher on a test over its content if (a) they are instructed prior to the message that they will be tested, and (b) if they are instructed that their test scores will apply toward a course grade.

The two conditions listed in the second hypothesis were described in this study as extrinsic test incentives. The results of the analysis of variance of the listening achievement test scores showed that the overall main effect due to incentives was significant. Multiple comparisons among the means of the three incentive conditions showed that the subjects who were warned that they would receive a test and would also receive a grade based on their

performance scored significantly higher than subjects who were either not aware they would be tested or who were warned of the test but were told that they would not be graded for their performance on the test. The mean scores of the latter two groups were not significantly different from one another. Therefore, the null hypothesis regarding the effects of awareness versus no awareness of the test situation could not be rejected, thus implying that awareness of the test situation is not sufficient incentive by itself to affect listening achievement. The null hypothesis regarding the effects of instructions informing the listeners that they would be graded on their test achievement scores was rejected, thus implying that being in a test situation is motivating to listeners only when some extra incentive such as a grade is attached to the test.

Knower, Phillips, and Koeppel (1945) reported that subjects who were aware they would be tested achieved somewhat higher scores on a listening achievement test than subjects who were unaware they would be tested. However, since their results did not reach statistical significance when judged by modern conventional levels of probability, the results of the present study concerning the nonsignificant incentive effect of warning listeners that they would be tested were consistent with the findings of the earlier study.

The finding of the present study that course grades provided an incentive which significantly increased listening achievement scores was consistent with the results of experiments by Bohn & Frandsen (1964) and Goodyear (1969). The present results

also clarify the findings of Bohn & Frandsen by separating the incentive variables into two types; awareness versus no awareness of the test situation and grade incentives versus no grade incentives. These two variables were confounded in the study by Bohn & Frandsen. By showing that awareness of the test situation did not significantly increase listening test scores, the present study confirmed the claim of the earlier study that the significant increase in listening achievement scores was attributable solely to the grade incentives. Also, since previous studies used much shorter messages than the present study, the conclusion that course grades provide a significant incentive to listening achievement can probably be generalized to conditions employing either brief or lengthy messages.

The third hypothesis stated,

> There will be an ordinal interaction between expressed interest and the levels of test incentive. Specifically, increasing the level of test incentive will decrease the difference in test scores between subjects expressing high versus low interest in the message.

The results of the analysis of variance of the listening achievement scores showed no significant interaction between interest and incentives. Thus, the null hypothesis regarding the interaction could not be rejected. Listeners who perceived the message as interesting had significantly higher achievement test scores across all three levels of incentive than listeners who perceived the message as uninteresting.

This finding does not support the claim that listening under test conditions differs from listening under normal conditions. Kelly (1963) argued that listening tests are not representative of

normal listening behavior because they only measure listening ability, thus failing to measure variables governed by what Weaver (1972b) called the willingness to listen. The key implication of the present study is that test incentives serve to increase listening achievement scores of subjects but do not negate or interact with perceived interest, a variable related to listeners' willingness to listen to a message. Therefore, listening achievement scores obtained under test conditions may be more representative of normal listening behavior than was previously believed.

REFERENCES

Allen, W. Readability of instructional film commentary. Journal of Applied Psychology, 1952, 36, 164-168

Armstrong, O. K. Beware the commercial faith healers, Reader's Digest, 1971, 100, 179-186.

Bohn, T. W., & Frandsen, K. D. Immediate and delayed recall as a function of mode of presentation and anticipation of reward. Central States Speech Journal, 1964, 15, 291-293.

Brandon, J. R. An experimental television study: The relative effectiveness of presenting factual information by the lecture, interview, and discussion methods. Speech Monographs, 1956, 23, 272-283.

Brown, J. I. Evaluating student performance in listening. Education, 1955, 75, 316-321.

Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Cartier, F. A. Listenability and "Human Interest." Speech Monographs, 1955, 22, 53-57.

Davis, R. E. Billy Sunday: Preacher-showman. The Southern Speech Journal, 1966, 32, 83-97.

Goodyear, F. H. An experimental study of the motivational effect of punishment and reward anticipation on the listening comprehension of college students. Unpublished Doctoral Dissertation, University of Texas at Austin, 1969.

Heath, M. A. An experimental analysis of the relations between interest, educability, and a score on an objective examination over the factual content of an informative speech. Unpublished Master's Thesis, Florida State University, 1951.

Karraker, M. E. An evaluation of the influence of interest and "set" on listening effectiveness in the basic communication class. Unpublished Doctoral Dissertation, University of Denver, 1951.

Kelly, C. M. Mental ability and personality factors in listening. Quarterly Journal of Speech, 1963, 49, 152-156.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole, 1968.

Klein, H. A. Interest and comprehension in sex-typed materials. Paper presented at the International Reading Association Conference, Kansas City, Missouri, May 1969, Educational Resources Information Center, Document ED 030 551.

Knower, F. H., Phillips, D. & Koeppel, F. Studies in listening to informative speaking. The Journal of Abnormal and Social Psychology, 1945, 40, 82-88.

Kramer, C. Y. Extension of multiple range tests to group means with unequal numbers of replications. Biometrics, 1956, 12, 307-310.

Livingston, H. M. An experimental study of effects of interest and authority upon understanding of broadcast information. Unpublished Doctoral Dissertation, University of Southern California, 1961.

Marten, M. E. The relationship between expressed interests and listening skills of children in the sixth grade. Unpublished Doctoral Dissertation, Indiana University, 1958.

Matter, G. A. The relation of auditors' expressed interest and ability to predict subsequent material. Unpublished Doctoral Dissertation, University of Pittsburgh, 1968.

Sewell, E. H. The effect of orienting instructions and monetary incentives on listening comprehension, Unpublished Doctoral Dissertation, Ohio University, 1972.

Spearritt, D. Listening comprehension--a factorial analysis. Melbourne: Australian Council for Educational Research, Series No. 76, 1962.

Trenaman, J. M. Understanding radio talks. Quarterly Journal of Speech, 1951, 37, 173-178.

Trenaman, J. M. Communication and comprehension. E. M. Hutchinson (Ed.), Great Britain: Longmans, Ltd., 1967.

Vernon, P. E. An investigation into the intelligibility of educational broadcasts. London: The British Broadcasting Corporation, Audience Research Department, Document No. 2328, 1950.

Weaver, C. H. An evaluation of a method of teaching listening improvement. Project No. 1-E-099, Grant No. OEG-5-71-0044(509), U. S. Department of Health, Education, and Welfare, 1972a.

Weaver, C. H. Human listening: Processes and behavior. Indianapolis: Bobbs-Merrill, 1972b.

Young, R. Q. An experimental investigation of reading and listening comprehension and of the use of readability formulas as measures of listenability. Unpublished Doctoral Dissertation, Ohio University, 1972.